AIL1020 Foundations of Statistics & Probability

Module 03

Bayes' Theorem

Dr. Rajlaxmi Chouhan
Associate Professor, Department of Electrical Engineering
IIT Jodhpur

Powered by Futurense

Learning Outcomes

Powered by

Futurense

By the end of this lesson, you will be able to

explain and apply conditional probability and the Bayes theorem to calculate the overall probability of an event based on different conditions or cases.

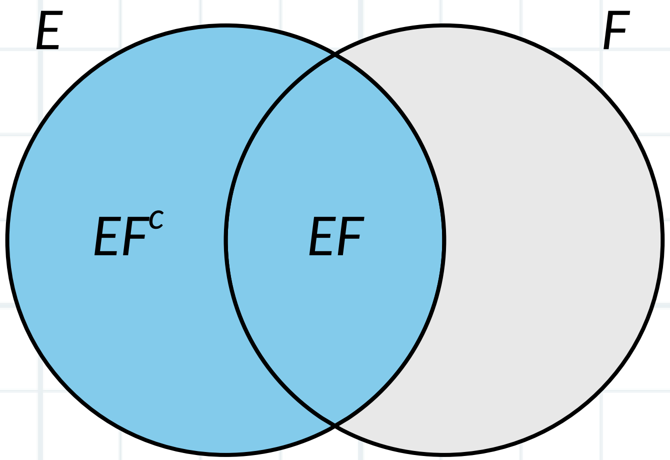
Conditioning and Weighted Average

Let *E* and *F* be events.

We may express *E* as

$$E = EF \cup EF^c$$

for, in order for a point to be in *E*, it must either be in both *E* and *F* or be in *E* but not in *F*.



As *EF* and EF^c are clearly mutually exclusive,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)[1 - P(F)] \end{aligned}$$

The probability of the event *E* is a **weighted average of the conditional probability of *E* given that *F* has occurred and the conditional probability of *E* given that *F* has not occurred**, with each conditional probability being given as much weight as the event it is conditioned on has of occurring.

Bayes Theorem

For two events A and B, where $P(B) > 0$:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Proof

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1)$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad (2)$$

$P(A \cap B) = P(B \cap A)$ — because intersection is commutative.

$$\Rightarrow \quad P(A \cap B) = P(B|A) \cdot P(A) \quad \Longrightarrow \quad \boxed{P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}}$$

Bayes' Theorem

Bayes' theorem is a fundamental formula in probability theory and statistics that describes how to update our belief in a hypothesis after observing new evidence.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

It provides a relationship between four key terms:

Prior probability of *A*	$P(A)$	Our belief in *A* before seeing evidence *B*
Likelihood	$P(B\|A)$	Likelihood of observing *B* given that *A* is true.
Evidence (or marginal probability)	$P(B)$	The overall probability of observing *B* under all scenarios – essentially a normalizing factor
Posterior Probability	$P(A\|B)$	Our updated belief in *A* after observing *B*

Law of Total Probability

We are calculating the **marginal probability** of an event (let's say event A) that can happen under **several different conditions**.

Let's assume we have:

A set of **mutually exclusive and exhaustive** events $B_1, B_2, \dots, B_n$ (i.e., exactly one of these B_i must be true)

The **Law of Total Probability** says:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

This is **conditioning on a partition** of the sample space.

Example

An insurance company believes that people can be divided into two classes:

If we assume that **30 % of the population is accident prone**, what is the probability that a new policy holder will have an accident within a year of purchasing a policy?

Calculate by first conditioning on whether or not the policy holder is accident prone.

A_1 : Event that the policy holder will have an accident within a year of purchase;

A : Event that the policy holder is accident prone

Accident prone

0.4

Not accident prone

0.2

Probability of having an accident at some time within a fixed 1-year period

$$P(A_1) = P(A_1|A)P(A) + P(A_1|A^c)P(A^c) = (.4)(.3) + (.2)(.7) = .26$$

Example

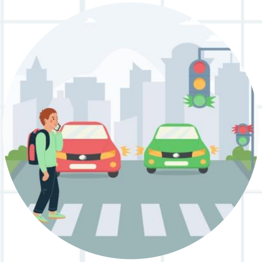
Suppose that a new policy holder has an accident within a year of purchasing his policy.

What is the probability that he is accident prone?

Initially, at the moment when the policy holder purchased his policy, we assumed there was a 30% chance that he was accident prone.

i.e. $P(A) = 0.3$

However, based on the fact that he has had an accident within a year, **we now reevaluate his probability of being accident prone as follows**.



Accident prone
0.4

Not accident prone
0.2

Probability of having an accident at some time within a fixed 1-year period

$$P(A|A_1) = \frac{P(AA_1)}{P(A_1)}$$

$$= \frac{P(A)P(A_1|A)}{P(A_1)} = \frac{(.3)(.4)}{.26} = \frac{6}{13} = .4615$$

Example

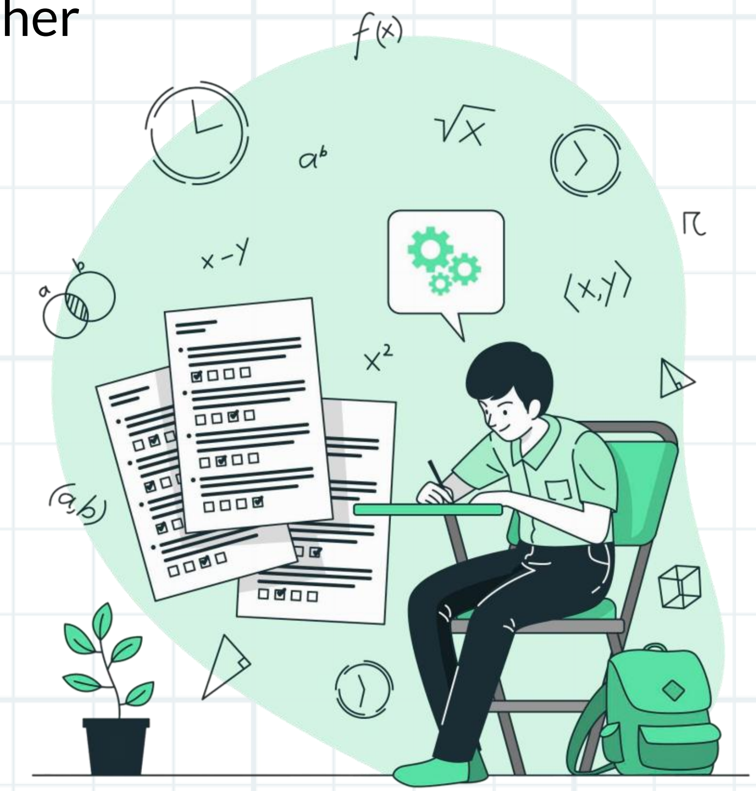
In answering a question on a multiple-choice test, a student either

Knows the answer	**Guesses** the answer
p	$1 - p$

Assume that a student who guesses at the answer will be correct with probability $1/m$.

where m is the number of multiple-choice alternatives.

What is the conditional probability that a student knew the answer to a question given that he answered it correctly?



Example

In answering a question on a multiple-choice test, a student either

Knows the answer

p $P(K)$

Guesses the answer

$1 - p$

Assume that a student who guesses at the answer will be correct with probability $1/m$.

where m is the number of multiple-choice alternatives.

What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

Let

C : Event that the student answers correctly

K : Event that the students knows the answer

We need to find

$P(K|C)$

$P(C|K) = P(CK)/P(K)$

$1 = P(KC)/p$

$P(KC) = p$

Example

In answering a question on a multiple-choice test, a student either

Knows the answer $P(K)$ p **Guesses** the answer $1 - p$

$P(K^C) = p$

What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

Let

C : Event that the student answers correctly

K : Event that the students knows the answer

We need to find

$P(K|C)$

To compute the probability that the student answers correctly, we condition on whether or not he knows the answer.

$$P(C) = P(C|K)P(K) + P(C|K^c)P(K^c)$$
$$= p + (1/m)(1-p)$$

$$\Rightarrow P(K|C) = \frac{p}{p + (1/m)(1-p)} = \frac{mp}{1 + (m-1)p}$$

Example

A laboratory blood test is **99 %** effective in detecting a certain disease when it is, in fact, present.

However, the test also yields a “false positive” result for **1 %** of the healthy persons tested.

If **.5 %** of the population actually has the disease,

what is the probability a person has the disease given that his test result is positive?

D: Event that the tested person has the disease

E: Event that his test results are positive

$P(D \mid E) = ?$

S. M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, 6th ed., 2021, Elsevier

Powered by Futurense

Example

A laboratory blood test is **99 %** effective in detecting a certain disease when it is, in fact, present.

However, the test also yields a "false positive" result for **1 %** of the healthy persons tested.

If **.5 %** of the population actually has the disease,

what is the probability a person has the disease given that his test result is positive?

D: Event that the tested person has the disease

E: Event that his test results are positive

***P*(*D* | *E*) = ?**

$$
\begin{aligned}
P(D|E) &= \frac{P(DE)}{P(E)} \\
&= \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\
&= \frac{(.99)(.005)}{(.99)(.005) + (.01)(.995)} \\
&= .3322
\end{aligned}
$$

Thus, only 33 % of those persons whose test results are positive actually have the disease!!!

S. M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, 6th ed., 2021, Elsevier

Example

Context

0.5% of population has the disease

Out of **200 people** tested:

- **0.5% of 200 = 1 person** actually **has** the disease
- **199 people** do **not** have the disease

$$P\,(\textit{Has disease} \mid \textit{Positive}) = \frac{\text{True positives}}{\text{Total positives}}$$

$$= 1/3$$

$$P(D \mid E) = 0.3322$$

Test

True Positives (test correctly identifies disease)

Among the 1 person **with the disease**:

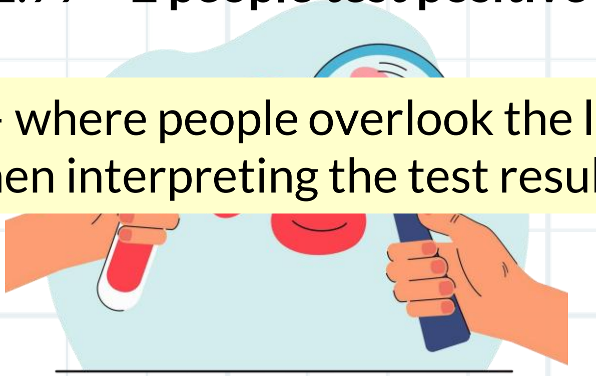
Test catches it correctly 99% of the time, so
0.99×1 **= 0.99 people test positive ~ 1 person**

False Positives (test wrongly says disease is present)

Among the 199 healthy people

Test wrongly says positive in 1% of cases, so
$0.01 \times 199 = 1.99$ ~ **2 people test positive falsely**

This is a classic example of **base rate fallacy** — where people overlook the low **prior probability** (base rate) of the disease when interpreting the test result.



S. M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, 6th ed., 2021, Elsevier

Summary

Law of total probability and Bayes theorem

Examples of using Bayes theorem in computing & updating posterior probability in the light of new evidence.

In the next video

More problems using Bayes' Theorem

Powered by Futurense

Thank you.